Brainomics: A management system for exploring and merging heterogeneous brain mapping data

Vincent Michel, Yannick Schwartz, Philippe Pinel, Olivier Cayrol, Antonio Moreno, Jean-Baptiste Poline, Vincent Frouin, Dimitri Papadopoulos-Orfanos

Brainomics - *A management system for exploring and merging heterogeneous brain mapping data*

V. Michel[1], Y. Schwartz[2,4], P. Pinel[3], O. Cayrol[1], A. Moreno[3], JB. Poline[4], V. Frouin[4], D. Papadopoulos Orfanos[4]

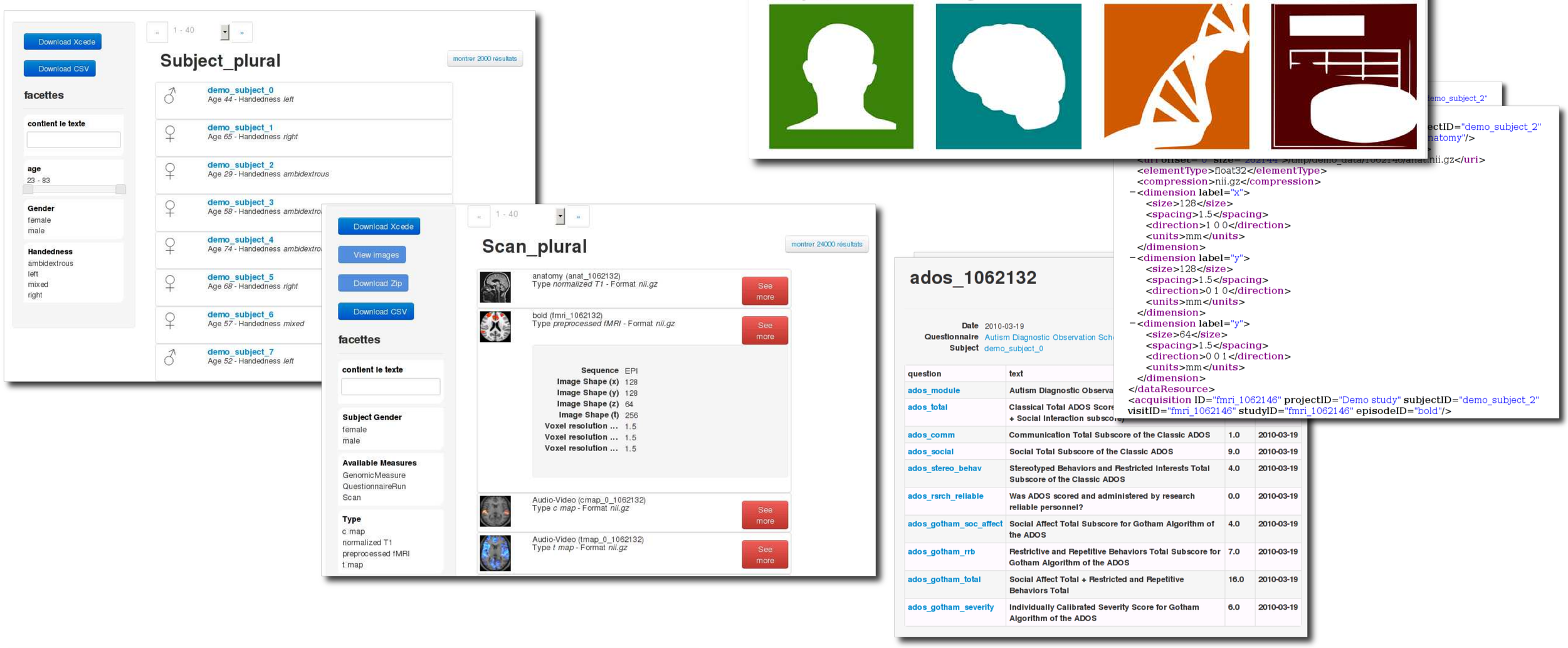
[1]Logilab, Paris, France, [2]INRIA Saclay Parietal Team, Paris, France, [3]Unicog, INSERM-CEA, Neurospin, Gif-sur-Yvette, France,[4]CEA, Neurospin, Gif-sur-Yvette, France

Introduction

- *Number of large datasets for brain mapping have been released [1, 2].*
- *Neuroimaging datasets more routinely include clinical data or genetics data.*
- *Exploitation requires*
  - ➔ *An efficient organization to integrate all the measures*
  - ➔ *An easy access to the relevant information.*
- *Neuroimaging [3] and genomics[4] databases are dedicated to their own field of research.*

http://www.brainomics.net/demo/

- Brings together **brain imaging and genetics data.**
- Relies on a **high-level query language (RQL).**
- Solution based on **CubicWeb,** a semantic framework.
- **Visualizing / exporting data** in several formats.



CubicWeb

http://www.cubicweb.org/

- ✔ **Data management framework,** 10 years of industrial uses (e.g. [5]).
- ✔ Well established core technologies: **SQL, Python, HTML5, Javascript.**
- ✔ Licensed under **LGPL** since 2008.
- ✔ Used in production environments since 2005.
- ✔ Fine-grained **security system** coupled to the data model definition.
- ✔ **Migration mechanisms** controls model version / ensures data integrity.

Data model

- ✔ Described in Python, using reusable modules called “cubes”.
- ✔ Modelisation of Scans, Questionnaires, Genomics results, Behavioral results, Subjects and Studies information.
- ✔ Data model optimized for large volumes (> 2000 subjects).
- ✔ Tested with several publicly available datasets [1, 2].

Query using RQL

- ✔ Similar to the W3C's SPARQL [6].
- ✔ Supports the basic operations (selection, insertion, etc.).
- ✔ Subquerying, ordering, couting, ...

Query all the female subjects of the database, with an age greater than 30

➔ *Any S WHERE S is Subject, S age > 30, S gender "female"*

Query all the Cmap scans of subjects with an age greater than 25, and that have a score greater than 4.0 for the "algebre" question

➔ *Any SA WHERE S is Subject, S age > 25, X is QuestionnaireRun, X concerns S, A is Answer, A questionnaire_run X, A question Q, Q text "algebre", A value > 4, SA is Scan, SA concerns S, SA type "c map"*

Views

- ✔Each query result can be seen using different views.
- ✔HTML pages, ZIP files, spreadsheets, XCEDE XML, ...
- ✔May include processing (stat. maps computed on the fly).

Conclusion

Contact: brainomics@logilab.fr

- ✔ *Open source solution to manage brain imaging datasets and associated meta data.*
- ✔ *Powerful querying and reporting tool, customized for emerging imaging-genetics field.*

This work was supported by grants from the French National Reseach Agency (ANR GENIM; ANR-10-BLAN-0128) and (ANR IA BRAINOMICS; ANR-10-BINF-04).

■[1] http://openfmri.org/data-sets
■[2] fcon_1000.projects.nitrc.org/indi/abide/
■[3] Olsen, M.D. (2007). The extensible neuroimaging archive toolkit. Neuroinformatics 5, 11–33
■[4] Vallon-Christersson, J. (2009). ‘BASE - 2nd generation software for microarray data management and analysis’
■[5] Publishing bibliographic records on the Web of data: opportunities for the BnF (French national Library). ESWC 2013
■[6] http://www.w3.org/TR/rdf-sparql-query/